

(19)



(11)

EP 1 209 662 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
07.02.2007 Bulletin 2007/06

(51) Int Cl.:
G10L 15/28 (2006.01)

(21) Application number: **01309945.2**

(22) Date of filing: **27.11.2001**

(54) Client-server based speech recognition

Client/Server basiertes Spracherkennungssystem

Reconnaissance de la parole en mode client-serveur

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR**

(30) Priority: **27.11.2000 JP 2000360203**

(43) Date of publication of application:
29.05.2002 Bulletin 2002/22

(73) Proprietor: **CANON KABUSHIKI KAISHA
Tokyo (JP)**

(72) Inventors:
• **Kushida, Akihiro,**
c/o Canon Kabushiki Kaisha
Ohta-ku,
Tokyo (JP)

• **Kosaka, Tetsuo,**
c/o Canon Kabushiki Kaisha
Ohta-ku,
Tokyo (JP)

(74) Representative: **Beresford, Keith Denis Lewis et al**
BERESFORD & Co.
16 High Holborn
London WC1V 6BX (GB)

(56) References cited:
EP-A- 0 945 851 **WO-A-99/26233**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description**FIELD OF THE INVENTION**

[0001] The present invention relates to a client-server speech recognition system for recognizing speech input at a client by a server, a speech recognition server, a speech recognition client, their control method, and a computer readable memory.

BACKGROUND OF THE INVENTION

[0002] In recent years, speech is used as an input interface in addition to a keyboard, mouse, and the like.

[0003] However, the recognition rate of speech recognition that recognizes input speech lowers and requires a longer processing time as the number of recognition words which are to undergo speech recognition becomes larger. For this reason, in an actual method, a plurality of recognition dictionaries or lexicons that register recognition words (e.g., pronunciations and notations) which are to undergo speech recognition are prepared, and are selectively used (a plurality of recognition dictionaries may be used at the same time).

[0004] Also, unregistered words cannot be recognized. As one of methods for solving this problem, a user dictionary or lexicon (prepared by the user to register recognition words which are to undergo speech recognition) may be used.

[0005] On the other hand, a client-server speech recognition system has been studied to implement speech recognition on a terminal with insufficient resources.

[0006] These three techniques are known to those who are skilled in the art, but a system that combines these three techniques has not been realized yet.

[0007] WO-A-99/26233 discloses a client-server speech recognition system in which speech data generated by a client is transmitted to a speech recognition server where the speech is recognised and the recognised speech is returned to the client.

SUMMARY OF THE INVENTION

[0008] According to one aspect, the present invention provides a client-server speech recognition system for recognizing speech input at a client by a server, the client comprising:

speech input means for inputting speech;

user dictionary holding means for holding a user dictionary formed by registering target recognition words designated by the user; and

transmission means for transmitting speech data input by said speech input means, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server, and

the server comprising:

recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields; determination means for determining one or more recognition dictionaries corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries and the user dictionary received from the client; and recognition means for recognizing the speech data using at least the recognition dictionary determined by said determination means.

[0009] Other features and advantages of the present invention will be apparent from the following description taken in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

BRIEF DESCRIPTION OF THE DRAWINGS**[0010]**

Fig. 1 is a block diagram showing the hardware arrangement of a speech recognition system of the first embodiment;

Fig. 2 is a block diagram showing the functional arrangement of the speech recognition system of the first embodiment;

Fig. 3 shows the configuration of a user dictionary of the first embodiment;

Fig. 4 shows a speech input window of the first embodiment;

Fig. 5 shows an identifier table of the first embodiment;

Fig. 6 is a flow chart showing the process executed by the speech recognition system of the first embodiment;

Fig. 7 shows the configuration of a user dictionary appended with input form identifiers according to the third embodiment; and

Fig. 8 shows the configuration of a user dictionary appended with recognition dictionary identifiers according to the third embodiment.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0011] Preferred embodiments of the present invention will be described in detail below with reference to the accompanying drawings.

[First Embodiment]

[0012] Fig. 1 shows the hardware arrangement of a speech recognition system of the first embodiment.

[0013] A CPU 101 systematically controls an entire cli-

ent 100. The CPU 101 loads programs stored in a ROM 102 onto a RAM 103, and executes various processes on the basis of the loaded programs. The ROM 102 stores various programs of processes to be executed by the CPU 101. The RAM 103 provides a storage area required to execute various programs stored in the ROM 102.

[0014] A secondary storage device 104 stores an OS and various programs. When the client 100 is implemented using not a general-purpose apparatus such as a personal computer or the like but a dedicated apparatus, the ROM 102 may store the OS and various programs. By loading the stored programs onto the RAM 103, the CPU 101 can execute processes. As the secondary storage device 104, a hard disk device, floppy disk drive, CD-ROM, or the like may be used. That is, storage media are not particularly limited.

[0015] A network I/F (interface) 105 is connected to a network I/F 205 of a server 200.

[0016] An input device 106 comprises a mouse, keyboard, microphone, and the like to allow input of various instructions to processes to be executed by the CPU 101, and can be used by simultaneously connecting these plurality of devices. An output device 107 comprises a display (CRT, LCD, or the like), and displays information input by the input device 106, and display windows which are controlled by various processes executed by the CPU 101. A bus 108 interconnects various building components of the client 100.

[0017] A CPU 201 systematically controls the entire server 200. The CPU 201 loads programs stored in a ROM 202 onto a RAM 203, and executes various processes on the basis of the loaded programs. The ROM 202 stores various programs of processes to be executed by the CPU 201. The RAM 203 provides a storage area required to execute various programs stored in the ROM 202.

[0018] A secondary storage device 204 stores an OS and various programs. When the server 200 is implemented using not a versatile apparatus such as a personal computer or the like but a dedicated apparatus, the ROM 202 may store the OS and various programs. By loading the stored programs onto the RAM 203, the CPU 201 can execute processes. As the secondary storage device 204, a hard disk device, floppy disk drive, CD-ROM, or the like may be used. That is, storage media are not particularly limited.

[0019] The network I/F 205 is connected to the network I/F 105 of the client 100. A bus 206 interconnects various building components of the server 200.

[0020] The functional arrangement of the speech recognition system of the first embodiment will be described below using Fig. 2.

[0021] Fig. 2 is a block diagram showing the functional arrangement of the speech recognition system of the first embodiment.

[0022] In the client 100, a speech input module 121 inputs speech uttered by the user via a microphone (input device 106), and A/D-converts input speech data

(speech recognition data) which is to undergo speech recognition. A communication module 122 sends a user dictionary 124a, speech recognition data 124b, dictionary management information 124c, and the like to the server 200. Also, the communication module 122 receives a speech recognition result of the sent speech recognition data 124b and the like from the server 200.

[0023] A display module 123 displays the speech recognition result received from the server 200 while storing it in, e.g., an input form which is displayed on the output device 107 by the process executed by the speech recognition system of this embodiment.

[0024] In the server 200, a communication module 221 receives the user dictionary 124a, speech recognition data 124b, dictionary management information 124c, and the like from the client 100. Also, the communication module 221 sends the speech recognition result of the speech recognition data 124b and the like to the client 100.

[0025] A dictionary management module 223 switches and selects a plurality of kinds of recognition dictionaries 225 (recognition dictionary 1 to recognition dictionary N, N: a positive integer) prepared for respective recognition fields (e.g., for names, addresses, alphanumeric symbols, and the like), and the user dictionary 124a received from the client 100 (may simultaneously use a plurality of kinds of dictionaries).

[0026] Note that the plurality of kinds of recognition dictionaries 225 are prepared for each dictionary management information 124c (input form identifier; to be described later) sent from the client 100. Each recognition dictionary 225 is appended with a recognition dictionary identifier indicating the recognition field of that recognition dictionary. The dictionary management module 223 manages an identifier table 223a that stores the recognition dictionary identifiers and input form identifiers in correspondence with each other, as shown in Fig. 5.

[0027] A speech recognition module 224 executes speech recognition using the recognition dictionary or dictionaries 225 and user dictionary 124a designated for speech recognition by the dictionary management module 223 on the basis of the speech recognition data 124b and dictionary management information 124c received from the client 100.

[0028] Note that the user dictionary 124a is prepared by the user to register recognition words which are to undergo speech recognition, and stores pronunciations and notations of words to be recognized in correspondence with each other, as shown in, e.g., Fig. 3.

[0029] The speech recognition data 124b may be either speech data A/D-converted by the speech input module 121 or data obtained by encoding that speech data.

[0030] The dictionary management information 124c indicates an input object and the like. For example, the dictionary management information 124c is an identifier (input form identifier) indicating the type of input form when the server 200 recognizes input speech and inputs

text data corresponding to that speech recognition result to each input form, which defines a speech input window displayed by the speech recognition system of the first embodiment, as shown in Fig. 4. The client 100 sends this input form identifier to the server 200 as the dictionary management information 124c. In the server 200, the dictionary management module 223 looks up the identifier table 223a to acquire a recognition dictionary identifier corresponding to the received input form identifier, and determines a recognition dictionary 225 to be used in speech recognition.

[0031] The process executed by the speech recognition system of the first embodiment will be explained below using Fig. 6.

[0032] Fig. 6 is a flow chart showing the process executed by the speech recognition system of the first embodiment.

[0033] In step S101, the client 100 sends the user dictionary 124a to the server 200.

[0034] In step S201, the server 200 receives the user dictionary 124a from the client 100.

[0035] In step S102, when speech is input to an input form as a target speech input, the client 100 sends the input form identifier of that input form to the server 200 as the dictionary management information 124c.

[0036] In step S202, the server 200 receives the input form identifier from the client 100 as the dictionary management information 124c.

[0037] In step S203, the server 200 looks up the identifier table 223a using the dictionary management information 124c to acquire a recognition dictionary identifier corresponding to the received input form identifier, and determines a recognition dictionary 225 to be used in speech recognition.

[0038] In step S103, the client 100 sends speech recognition data 124b, which is speech input as text data to be input to each input form, to the server 200.

[0039] In step S204, the server 200 receives the speech recognition data corresponding to each input form from the client 100.

[0040] In step S205, the server 200 executes speech recognition of the speech recognition data 124b in the speech recognition module 224 using the recognition dictionary 225 and user dictionary 124a designated for speech recognition by the dictionary management module 223.

[0041] In the first embodiment, all recognition words contained in the user dictionary 124a sent from the client 100 to the server 200 are used in speech recognition by the speech recognition module 224.

[0042] In step S206, the server 200 sends the speech recognition result obtained by the speech recognition module 224 to the client 100.

[0043] In step S104, the client 100 receives the speech recognition result corresponding to each input form from the server 200, and stores text data corresponding to the speech recognition result in the corresponding input form.

[0044] The client 100 checks in step S105 if the processing is to be ended. If the processing is not to be ended (NO in step S105), the flow returns to step S102 to repeat the processing. On the other hand, if the processing is to be ended (YES in step S105), the client 100 informs the server 200 of end of the processing, and ends the processing.

[0045] It is checked in step S207 if a processing end instruction from the client 100 is detected. If no processing end instruction is detected (NO in step S207), the flow returns to step S202 to repeat the above processes. On the other hand, if the processing end instruction is detected (YES in step S207), the processing ends.

[0046] In the above processing, when speech is input to an input form as a target speech input, the dictionary management information 124c corresponding to that input form is sent from the client 100 to the server 200. Alternatively, the dictionary management information 124c may be sent when the input form as a target speech input is focused by an instruction from the input device 106 (the input form as a target speech input is determined).

[0047] In the server 200, speech recognition is made after all speech recognition data 124b are received. Alternatively, every time speech is input as text data to a given input form, that the portion of speech recognition data 124b may be sent to the server 200 frame by frame (for example, one frame is 10 msec speech data), and speech recognition may be made in real time.

[0048] As described above, according to the first embodiment, in the client-server speech recognition system, since the server 200 executes speech recognition of speech recognition data 124b using both an appropriate recognition dictionary 225 and the user dictionary 124a, the speech recognition precision in the server 200 can be improved while reducing the processing load and use of storage resources associated with speech recognition in the client 100.

[Second Embodiment]

[0049] In the first embodiment, if no recognition words to be stored in the user dictionary 124a are generated, since the user dictionary 124a need not be used, the server 200 may use all recognition words in the user dictionary 124a in recognition only when a use request of the user dictionary 124a is received from the client 100.

[0050] In this case, a flag indicating if the user dictionary 124a is used is added as the dictionary management information 124c, thus informing the server 200 of the presence/absence of use of the user dictionary 124a.

[Third Embodiment]

[0051] Since some target recognition words in the user dictionary 124a are not used depending on an input object, situation, and the like, only specific recognition words in the user dictionary 124a may be used in recog-

dition depending on the input object and situation.

[0052] In such case, when the user dictionary is managed by designating input form identifiers for respective recognition words, as shown in Fig. 7, only recognition words having an input form identifier of the input form used in speech input can be used in recognition. Alternatively, a plurality of input form identifiers may be designated for a given recognition word. In addition, the user dictionary may be managed by designating recognition dictionary identifiers in place of input form identifiers, as shown in Fig. 8.

[Fourth Embodiment]

[0053] By combining the second and third embodiments, the efficiency of the speech recognition process of the speech recognition module 224 can be further improved.

[Fifth Embodiment]

[0054] Most of the processes of the apparatus of the present invention can be implemented by programs. As described above, since the apparatus can use a general-purpose apparatus such as a personal computer, the present invention is also achieved by supplying a storage medium, which records a program code of a software program that can implement the functions of the above-mentioned embodiments to a system or apparatus, and reading out and executing the program code stored in the storage medium by a computer of the system or apparatus. In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiments, and the storage medium which stores the program code constitutes the present invention. As the storage medium for supplying the program code, for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, magnetic tape, nonvolatile memory card, ROM, and the like may be used.

[0055] The present invention can also be achieved by supplying the storage medium that records the program code to a computer, and executing some or all of actual processes executed by an OS running on the computer. Furthermore, the functions of the above-mentioned embodiments may be implemented by some or all of actual processing operations executed by a CPU or the like arranged in a function extension board or a function extension unit, which is inserted in or connected to the computer, after the program code read out from the storage medium is written in a memory of the extension board or unit. When the present invention is applied to the storage medium, that storage medium stores a program code corresponding to the flow chart shown in Fig. 3.

[0056] As many apparently widely different embodiments of the present invention can be made without departing from the scope thereof, it is to be understood that the invention is not limited to the specific embodiments

thereof except as defined in the appended claims.

Claims

1. A client-server speech recognition system for recognizing speech input at a client (100) by a server (200), the client (100) comprising:

speech input means (121, 123) for inputting speech;

user dictionary holding means for holding a user dictionary (124a) formed by registering target recognition words designated by the user; and transmission means (122) for transmitting speech data (124b) input by said speech input means (121), dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary (124a) to the server (200), and

the server (200) comprising:

recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries (225) prepared for respective recognition fields;

determination means (223) for determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from the plurality of kinds of recognition dictionaries (225) and the user dictionary (124a) received from the client (100); and

recognition means (224) for recognizing the speech data (124b) using at least the recognition dictionary (225) determined by said determination means (223).

2. The system according to claim 1, wherein said recognition means (224) is operable to recognize the speech data (124b) using the recognition dictionary (225) determined by said determination means (223), and the user dictionary (124a) received from the client (100).
3. The system according to claim 1 or 2, wherein said speech input means (121, 123) comprises display means (123) for displaying an input form as a target speech input, and wherein the dictionary management information (124c) is an input form identifier that indicates a type of input form.
4. The system according to any of claims 1 to 3, wherein the dictionary management information (124c) contains information indicating if the user dictionary (124a) is to be used in recognition of the speech data

(124b).

5. The system according to any preceding claim, wherein the user dictionary (124a) is formed by storing pronunciations and notations of the target recognition words in correspondence with each other. 5
6. The system according to claim 3, wherein the user dictionary (124a) is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other. 10
7. The system according to any preceding claim, wherein the user dictionary (124a) is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words. 15
8. The system according to any preceding claim, wherein the speech data (124b) is data obtained by encoding that speech data. 20
9. A method of controlling a client-server speech recognition system for recognizing speech input at a client (100) by a server (200), comprising: 25
 - a speech input step of inputting speech; and
 - a transmission step of transmitting speech data (124b) input in the speech input step, dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) held in the client (100) which is formed by registering target recognition words designated by a user, to the server; 30
 - a determination step of determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from a plurality of kinds of recognition dictionaries (225) held in the server (200), which are prepared for respective recognition fields, and the user dictionary (124a) received from the client (100); and 40
 - a recognition step of recognizing the speech data (124b) using at least the recognition dictionary (225) determined in the determination step. 45
10. The method according to claim 9, wherein the recognition step includes a step of recognizing the speech data (124b) using the recognition dictionary (225) determined in the determination step, and the user dictionary (124a) received from the client (100). 50
11. The method according to claim 9 or 10, wherein the speech input step comprises a display step of displaying an input form as a target speech input, and wherein the dictionary management information 55

(124c) is an input form identifier that indicates a type of input form.

12. The method according to any of claims 9 to 11, wherein the dictionary management information (124c) contains information indicating if the user dictionary (124a) is to be used in recognition of the speech data (124b).
13. The method according to any of claims 9 to 12, wherein the user dictionary (124a) is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.
14. The method according to claim 11, wherein the user dictionary (124a) is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.
15. The method according to any of claims 9 to 14, wherein the user dictionary (124a) is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.
16. The method according to any of claims 9 to 15, wherein the speech data (124b) is data obtained by encoding that speech data.
17. A computer readable memory that stores a program code of control of a client-server speech recognition system for recognizing speech input at a client by a server, comprising:
 - a program code of a speech input step of inputting speech;
 - a program code of a transmission step of transmitting speech data (124b) input in the speech input step, dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) held in the client (100) which is formed by registering target recognition words designated by the user, to the server (200)
 - a program code of a determination step of determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from a plurality of kinds of recognition dictionaries (225) held in the server (200) which are prepared for respective recognition fields, and the user dictionary (124a) received from the client (100); and
 - a program code of a recognition step of recognizing the speech data (124b) using at least the recognition dictionary (225) determined in the

determination step.

- 18.** A speech recognition server (200) for recognizing speech input at a client (100), and sending a recognition result to the client (100), comprising:

reception means for receiving, from the client (100), speech data (124b), dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) formed by registering target recognition words designated by a user;
recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries (225) prepared for respective recognition fields;
determination means (223) for determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from the plurality of kinds of recognition dictionaries (225) and the user dictionary (124a) received from the client (100); and
recognition means (224) for recognizing the speech data using at least the recognition dictionary (225) determined by said determination means (223).

- 19.** The server according to claim 18, wherein said recognition means (224) is operable to recognize the speech data (124b) using the recognition dictionary (225) determined by said determination means (223), and the user dictionary (124a) received from the client (100).

- 20.** The server according to claim 18 or 19, wherein the speech data (124b) is data obtained by encoding that speech data.

- 21.** A speech recognition client (100) for sending input speech to be recognized to a server (200), and receiving a recognition result of that speech, comprising:

speech input means (121, 123) for inputting speech;
user dictionary holding means for holding a user dictionary (124a) formed by registering target recognition words designated by a user; and
transmission means (122) for transmitting speech data (124b) input by said speech input means, dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary (124a) to the server (200).

- 22.** The client according to claim 21, wherein said

speech input means (121, 123) comprises display means (123) for displaying an input form as a target speech input, and wherein the dictionary management information (124c) is an input form identifier that indicates a type of input form.

- 23.** The client according to claim 21 or 22, wherein the dictionary management information (124c) contains information indicating if the user dictionary (124a) is to be used in recognition of the speech data.

- 24.** The client according to any of claims 21 to 23, wherein the user dictionary (124a) is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.

- 25.** The client according to claim 22, wherein the user dictionary (124a) is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.

- 26.** The client according to any of claims 21 to 25, wherein the user dictionary (124a) is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.

- 27.** The client according to any of claims 21 to 25, wherein the speech data (124b) is data obtained by encoding that speech data.

- 28.** A method of controlling a speech recognition server (200) for recognizing speech input at a client (100), and sending a recognition result to the client (100), comprising:

a reception step of receiving, from the client, speech data (124b), dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) formed by registering target recognition words designated by a user;
a determination step of determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from a plurality of kinds of recognition dictionaries (225) held in the server (200), which are prepared for respective recognition fields, and the user dictionary (124a) received from the client (100); and
a recognition step of recognizing the speech data (124b) using at least the recognition dictionary (225) determined in the determination step.

- 29.** The method according to claim 28, wherein the recognition step includes a step of recognizing the

speech data (124b) using the recognition dictionary (225) determined in the determination step, and the user dictionary (124a) received from the client (100).

30. The method according to claim 28 or 29, wherein the speech data (124b) is data obtained by encoding that speech data. 5
31. A method of controlling a speech recognition client (100) for sending input speech to be recognized to a server (200), and receiving a recognition result of that speech, comprising: 10
a speech input step of inputting speech; and
a transmission step of transmitting speech data (124b) input in the speech input step, dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) held in the client (100) which is formed by registering target recognition words designated by a user, to the server (200). 15 20
32. The method according to claim 31, wherein the speech input step comprises a display step of displaying an input form as a target speech input, and wherein the dictionary management information (124c) is an input form identifier that indicates a type of input form. 25
33. The method according to claim 31 or 32, wherein the dictionary management information (124c) contains information indicating if the user dictionary (124a) is to be used in recognition of the speech data (124b). 30
34. The method according to any of claims 31 to 33, wherein the user dictionary (124a) is formed by storing pronunciations and notations of the target recognition words in correspondence with each other. 35 40
35. The method according to claim 32, wherein the user dictionary (124a) is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other. 45
36. The method according to any of claims 31 to 35, wherein the user dictionary (124a) is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words. 50
37. The method according to any of claims 31 to 36, wherein the speech data (124b) is data obtained by encoding that speech data. 55
38. A computer readable memory that stores a program code of control of a speech recognition server (200)

for recognizing speech input at a client (100), and sending a recognition result to the client (100), comprising:

- a program code of a reception step of receiving, from the client (100), speech data (124b), dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) formed by registering target recognition words designated by a user; a program code of a determination step of determining one or more recognition dictionaries (225) corresponding to the dictionary management information (124c) received from the client (100) from a plurality of kinds of recognition dictionaries (225) held in the server (200), which are prepared for respective recognition fields, and the user dictionary (124a) received from the client (100); and
a program code of a recognition step of recognizing the speech data (124b) using at least the recognition dictionary (225) determined in the determination step.
39. A computer readable memory that stores a program code of control of a speech recognition client (100) for sending input speech to be recognized to a server (200), and receiving a recognition result of that speech, comprising:
a program code of a speech input step of inputting speech; and
a program code of a transmission step of transmitting speech data (124b) input in the speech input step, dictionary management information (124c) used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary (124a) held in the client (100) which is formed by registering target recognition words designated by a user, to the server (200).
40. Processor implementable instructions product causing a programmable computer device to carry out all the steps of the method of any of claims 28 to 37, when the instructions product is run on said programmable computer device.

Patentansprüche

1. Client-/Server-Spracherkennungssystem zum Erkennen einer Spracheingabe bei einem Client (100) durch einen Server (200), wobei der Client (100) umfasst:
eine Spracheingabeeinrichtung (121, 123) zum

Eingeben von Sprache,
eine Benutzerwörterbuchhalteeinrichtung zum Halten eines Benutzerwörterbuchs (124a), das durch Registrieren von durch den Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, und

eine Übertragungseinrichtung (122) zum Übertragen von durch die Spracheingabeeinrichtung (121) eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet werden, und des Benutzerwörterbuchs (124a) zu dem Server (200), und

wobei der Server (200) umfasst:

eine Erkennungswörterbuchhalteeinrichtung zum Halten einer Vielzahl von Arten von Erkennungswörterbüchern (225), die für jeweilige Erkennungsfelder ausgebildet sind,
eine Bestimmungseinrichtung (223) zum Bestimmen von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen (124c) entsprechen, aus der Vielzahl von Arten von Erkennungswörterbüchern (225) und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und
eine Erkennungseinrichtung (224) zum Erkennen der Sprachdaten (124b) unter Verwendung zumindest des durch die Bestimmungseinrichtung (223) bestimmten Erkennungswörterbuchs (225).

2. System gemäß Anspruch 1, wobei die Erkennungseinrichtung (224) zum Erkennen der Sprachdaten (124b) unter Verwendung des durch die Bestimmungseinrichtung (223) bestimmten Erkennungswörterbuchs (225) und des von dem Client (100) empfangenen Benutzerwörterbuchs (124a) betreibbar ist.
3. System gemäß Anspruch 1 oder 2, wobei die Spracheingabeeinrichtung (121, 123) eine Anzeigeeinrichtung (123) zum Anzeigen einer Eingabemaske als eine Zielspracheingabe umfasst, und wobei die Wörterbuchverwaltungsinformationen (124c) eine Eingabenmaskenkennung darstellen, die eine Art der Eingabemaske angibt.
4. System gemäß einem der Ansprüche 1 bis 3, wobei die Wörterbuchverwaltungsinformationen (124c) Informationen enthalten, die angeben, ob das Benutzerwörterbuch (124a) bei einer Erkennung der Sprachdaten (124b) zu verwenden ist.

5. System gemäß einem der vorstehenden Ansprüche, wobei das Benutzerwörterbuch (124a) durch Speichern von Aussprachen und Schreibweisen der Zielerkennungswörter in Entsprechung miteinander ausgebildet ist.

6. System gemäß Anspruch 3, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Eingabemaskekennung und den Zielerkennungswörtern in Entsprechung miteinander ausgebildet ist.

7. System gemäß einem der vorstehenden Ansprüche, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Erkennungswörterbuchkennung, die Erkennungsfelder der Vielzahl von Arten von Erkennungswörterbüchern angibt, und der Zielerkennungswörter ausgebildet ist.

8. System gemäß einem der vorstehenden Ansprüche, wobei die Sprachdaten (124b) durch Codierung jener Sprachdaten erhaltene Daten sind.

9. Verfahren zum Steuern eines Client-/Server-Spracherkennungssystems zum Erkennen einer Spracheingabe bei einem Client (100) durch einen Server (200), mit:

einem Spracheingabeschritt des Eingebens von Sprache, und

einem Übertragungsschritt des Übertragens von in dem Spracheingabeschritt eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet werden, und eines in dem Client (100) gehaltenen Benutzerwörterbuchs (124a), das durch ein Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, zu dem Server, einem Bestimmungsschritt des Bestimmens von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen entsprechen, aus einer Vielzahl von Arten von in dem Server (200) gehaltenen Erkennungswörterbüchern (225), die für jeweilige Erkennungsfelder ausgebildet sind, und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und einem Erkennungsschritt des Erkennens der Sprachdaten (124b) unter Verwendung zumindest des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225).

10. Verfahren gemäß Anspruch 9, wobei der Erkennungsschritt einen Schritt des Erkennens der

Sprachdaten (124b) unter Verwendung des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225) und des von dem Client (100) empfangenen Benutzerwörterbuchs (124a) enthält.

11. Verfahren gemäß Anspruch 9 oder 10, wobei der Spracheingabeschritt einen Anzeigeschritt des Anzeigens einer Eingabemaske als eine Zielspracheingabe umfasst, und wobei die Wörterbuchverwaltungsinformationen (124c) eine Eingabemaskenkennung darstellen, die eine Art von Eingabemaske angibt. 5
12. Verfahren gemäß einem der Ansprüche 9 bis 11, wobei die Wörterbuchverwaltungsinformationen (124c) Informationen enthalten, die angeben, ob das Benutzerwörterbuch (124a) bei einer Erkennung der Sprachdaten (124b) zu verwenden ist. 10
13. Verfahren gemäß einem der Ansprüche 9 bis 12, wobei das Benutzerwörterbuch (124a) durch Speichern von Aussprachen und Schreibweisen der Zielerkennungswörter in Entsprechung miteinander ausgebildet wird. 15
14. Verfahren gemäß Anspruch 11, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Eingabemaskenkennung und der Zielerkennungswörter in Entsprechung miteinander ausgebildet wird. 20
15. Verfahren gemäß einem der Ansprüche 9 bis 14, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Erkennungswörterbuchkennung, die Erkennungsfelder der Vielzahl von Arten von Erkennungswörterbüchern angibt, und der Zielerkennungswörter ausgebildet wird. 25
16. Verfahren gemäß einem der Ansprüche 9 bis 15, wobei die Sprachdaten (124b) durch Codieren jener Sprachdaten erhaltene Daten sind. 30
17. Computerlesbarer Speicher, der einen Programmcode zur Steuerung eines Client-/Server-Spracherkennungssystems zum Erkennen einer Spracheingabe bei einem Client durch einen Server speichert, mit: 35

einem Programmcode eines Spracheingabeschritts des Eingebens von Sprache, 50
einem Programmcode eines Übertragungsschritts des Übertragens von in dem Spracheingabeschritt eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet werden, und eines in dem Client (100) ge- 55

haltenen Benutzerwörterbuchs (124a), das durch Registrieren von durch den Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, zu dem Server (200), einem Programmcode eines Bestimmungsschritts des Bestimmens von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen (124c) entsprechen, aus einer Vielzahl von Arten von in dem Server (200) gehaltenen Erkennungswörterbüchern (225), die für jeweilige Erkennungsfelder ausgebildet sind, und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und einem Programmcode eines Erkennungsschritts des Erkennens der Sprachdaten (124b) unter Verwendung zumindest des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225).

18. Spracherkennungsserver (200) zum Erkennen einer Spracheingabe bei einem Client (100) und zum Senden eines Erkennungsergebnisses zu dem Client (100), mit:

einer Empfangseinrichtung zum Empfangen von dem Client (100) von Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und eines Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, einer Erkennungswörterbuchhalteeinrichtung zum Halten einer Vielzahl von Arten von für jeweilige Erkennungsfelder ausgebildeten Erkennungswörterbüchern (225), einer Bestimmungseinrichtung (223) zum Bestimmen von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen (124c) entsprechen, aus der Vielzahl von Arten von Erkennungswörterbüchern (225) und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und einer Erkennungseinrichtung (224) zum Erkennen der Sprachdaten unter Verwendung zumindest des durch die Bestimmungseinrichtung (223) bestimmten Erkennungswörterbuchs (225).

19. Server gemäß Anspruch 18, wobei die Erkennungseinrichtung (224) zum Erkennen der Sprachdaten (124b) unter Verwendung des durch die Bestimmungseinrichtung (223) bestimmten Erkennungswörterbuchs (225) und des von dem Client (100)

empfangenen Benutzerwörterbuchs (124a) betreibbar ist.

20. Server gemäß Anspruch 18 oder 19, wobei die Sprachdaten (124b) durch Codieren jener Sprachdaten erhaltene Daten sind.

21. Spracherkennungsclient (100) zum Senden von durch einen Server (200) zu erkennender Eingabesprache und zum Empfangen eines Erkennungsergebnisses jener Sprache, mit:

einer Spracheingabeeinrichtung (121, 123) zum Eingeben von Sprache,
einer Benutzerwörterbuchhalteeinrichtung zum Halten eines Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, und
einer Übertragungseinrichtung (122) zum Übertragen von durch die Spracheingabeeinrichtung eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und des Benutzerwörterbuchs (124a) zu dem Server (200).

22. Client gemäß Anspruch 21, wobei die Spracheingabeeinrichtung (121, 123) eine Anzeigeeinrichtung (123) zum Anzeigen einer Eingabemaske als eine Zielspracheingabe umfasst, und wobei die Wörterbuchverwaltungsinformationen (124c) eine Eingabemaskenkennung darstellen, die eine Art von Eingabemaske angibt.

23. Client gemäß Anspruch 21 oder 22, wobei die Wörterbuchverwaltungsinformationen (124c) Informationen enthalten, die angeben, ob das Benutzerwörterbuch (124a) bei einer Erkennung der Sprachdaten zu verwenden ist.

24. Client gemäß einem der Ansprüche 21 bis 23, wobei das Benutzerwörterbuch (124a) durch Speichern von Aussprachen und Schreibweisen der Zielerkennungswörter in Entsprechung miteinander ausgebildet ist.

25. Client gemäß Anspruch 22, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern zumindest einer Eingabemaskenkennung und der Zielerkennungswörter in Entsprechung miteinander ausgebildet ist.

26. Client gemäß einem der Ansprüche 21 bis 25, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern zumindest einer Erkennungswörterbuchken-

nung, die Erkennungsfelder der Vielzahl von Arten von Erkennungswörterbüchern angibt, und der Zielerkennungswörter ausgebildet ist.

27. Client gemäß einem der Ansprüche 21 bis 25, wobei die Sprachdaten (124b) durch Codieren jener Sprachdaten erhaltene Daten sind.

28. Verfahren zum Steuern eines Spracherkennungsservers (200) zum Erkennen einer Spracheingabe bei einem Client (100) und zum Senden eines Erkennungsergebnisses zu dem Client (100), mit:

einem Empfangsschritt des Empfangens von dem Client von Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und eines Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist,

einem Bestimmungsschritt des Bestimmens von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen (124c) entsprechen, aus einer Vielzahl von Arten von in dem Server (200) gehaltenen Erkennungswörterbüchern (225), die für jeweilige Erkennungsfelder ausgebildet sind, und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und
einem Erkennungsschritt des Erkennens der Sprachdaten (124b) unter Verwendung zumindest des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225).

29. Verfahren gemäß Anspruch 28, wobei der Erkennungsschritt einen Schritt des Erkennens der Sprachdaten (124b) unter Verwendung des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225) und des von dem Client (100) empfangenen Benutzerwörterbuchs (124a) enthält.

30. Verfahren gemäß Anspruch 28 oder 29, wobei die Sprachdaten (124b) durch Codieren jener Sprachdaten erhaltene Daten sind.

31. Verfahren zum Steuern eines Spracherkennungsclients (100) zum Senden von durch einen Server (200) zu erkennender Eingabesprache und zum Empfangen eines Erkennungsergebnisses von jener Sprache, mit:

einem Spracheingabeschritt des Eingebens von Sprache, und
einem Übertragungsschritt des Übertragens

- von in dem Spracheingabeschritt eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und eines in dem Client (100) gehaltenen Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, zu dem Server (200).
32. Verfahren gemäß Anspruch 31, wobei der Spracheingabeschritt einen Anzeigeschritt des Anzeigens einer Eingabemaske als eine Zielspracheingabe umfasst, und wobei die Wörterbuchverwaltungsinformationen (124c) eine Eingabemaskenkennung darstellen, die eine Art von Eingabemaske angibt.
33. Verfahren gemäß Anspruch 31 oder 32, wobei die Wörterbuchverwaltungsinformationen (124c) Informationen enthalten, die angeben, ob das Benutzerwörterbuch (124a) bei einer Erkennung der Sprachdaten (124b) zu verwenden ist.
34. Verfahren gemäß einem der Ansprüche 31 bis 33, wobei das Benutzerwörterbuch (124a) durch Speichern von Aussprachen und Schreibweisen der Zielerkennungswörter in Entsprechung miteinander ausgebildet ist.
35. Verfahren gemäß Anspruch 32, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Eingabemaskenkennung und der Zielerkennungswörter in Entsprechung miteinander ausgebildet ist.
36. Verfahren gemäß einem der Ansprüche 31 bis 35, wobei das Benutzerwörterbuch (124a) ebenso durch Speichern von zumindest einer Erkennungswörterbuchkennung, die Erkennungsfelder der Vielzahl von Arten von Erkennungswörterbüchern angibt, und der Zielerkennungswörter ausgebildet ist.
37. Verfahren gemäß einem der Ansprüche 31 bis 36, wobei die Sprachdaten (124b) durch Codieren jener Sprachdaten erhaltene Daten sind.
38. Computerlesbarer Speicher, der einen Programmcode zur Steuerung eines Spracherkennungsservers (200) zum Erkennen einer Spracheingabe bei einem Client (100) und zum Senden eines Erkennungsergebnisses zu dem Client (100) speichert, mit:
- einem Programmcode eines Empfangsschritts des Empfangens von dem Client (100) von Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und eines Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, einem Programmcode eines Bestimmungsschritts des Bestimmens von einem oder mehreren Erkennungswörterbüchern (225), die den von dem Client (100) empfangenen Wörterbuchverwaltungsinformationen (124c) entsprechen, aus einer Vielzahl von Arten von in dem Server (200) gehaltenen Erkennungswörterbüchern (225), die für jeweilige Erkennungsfelder ausgebildet sind, und dem von dem Client (100) empfangenen Benutzerwörterbuch (124a), und einem Programmcode eines Erkennungsschritts des Erkennens der Sprachdaten (124b) unter Verwendung zumindest des in dem Bestimmungsschritt bestimmten Erkennungswörterbuchs (225).
39. Computerlesbarer Speicher, der einen Programmcode zur Steuerung eines Spracherkennungsclients (100) zum Senden einer durch einen Server (200) zu erkennenden Eingabesprache und zum Empfangen eines Erkennungsergebnisses jener Sprache speichert, mit:
- einem Programmcode eines Spracheingabeschritts des Eingebens von Sprache, und einem Programmcode eines Übertragungsschritts des Übertragens von in dem Spracheingabeschritt eingegebenen Sprachdaten (124b), von Wörterbuchverwaltungsinformationen (124c), die zum Bestimmen eines Erkennungsfeldes eines zum Erkennen der Sprachdaten verwendeten Erkennungswörterbuchs verwendet sind, und eines in dem Client (100) gehaltenen Benutzerwörterbuchs (124a), das durch Registrieren von durch einen Benutzer ausgewiesenen Zielerkennungswörtern ausgebildet ist, zu dem Server (200).
40. Prozessorimplementierbares Anweisungsprodukt, das eine programmierbare Computervorrichtung veranlasst, alle Schritte des Verfahrens gemäß einem der Ansprüche 28 bis 37 auszuführen, wenn das Anweisungsprodukt auf der programmierbaren Computervorrichtung läuft.

Revendications

1. Système client-serveur de reconnaissance de la parole pour la reconnaissance par un serveur (200) de la parole fournie en entrée sur un client (100), le client (100) comprenant :

- un moyen d'entrée de parole (121, 123) pour fournir en entrée de la parole ;
 un moyen de conservation de dictionnaire personnel pour conserver un dictionnaire personnel (124a) formé en enregistrant des mots de reconnaissance cibles désignés par l'utilisateur ; et
 un moyen de transmission (122) pour transmettre au serveur (200) des données de parole (124b) fournies en entrée par ledit moyen d'entrée de parole (121), des informations de gestion de dictionnaire (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et le dictionnaire personnel (124a), et
 le serveur (200) comprenant :
- un moyen de conservation de dictionnaire de reconnaissance pour conserver une pluralité de types de dictionnaires de reconnaissance (225) préparés pour des domaines de reconnaissance respectifs ;
 un moyen de détermination (223) pour déterminer un ou plusieurs dictionnaires de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) à partir de la pluralité de types de dictionnaires de reconnaissance (225) et du dictionnaire personnel (124a) reçu du client (100) ; et
 un moyen de reconnaissance (224) pour reconnaître les données de parole (124b) en utilisant au moins le dictionnaire de reconnaissance (225) déterminé par ledit moyen de détermination (223).
2. Système selon la revendication 1, dans lequel ledit moyen de reconnaissance (224) a pour fonction de reconnaître les données de parole (124b) en utilisant le dictionnaire de reconnaissance (225) déterminé par ledit moyen de détermination (223) et le dictionnaire personnel (124a) reçu du client (100).
 3. Système selon la revendication 1 ou 2, dans lequel ledit moyen d'entrée de parole (121, 123) comprend un moyen d'affichage (123) pour afficher une formule d'entrée en tant qu'entrée de parole cible, et dans lequel les informations de gestion de dictionnaires (124c) sont un identificateur de forme d'entrée qui indique un type de forme d'entrée.
 4. Système selon l'une quelconque des revendications 1 à 3, dans lequel les informations de gestion de dictionnaires (124c) contiennent des informations indiquant si le dictionnaire personnel (124a) doit être utilisé dans la reconnaissance des données de parole (124b).
 5. Système selon l'une quelconque des revendications précédentes, dans lequel le dictionnaire personnel (124a) est formé en stockant des prononciations et des notations des mots de reconnaissance cibles en correspondance les uns avec les autres.
 6. Système selon la revendication 3, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins un identificateur de forme d'entrée et les mots de reconnaissance cibles en correspondance les uns avec les autres.
 7. Système selon l'une quelconque des revendications précédentes, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins un des identificateurs de dictionnaires de reconnaissance indiquant des domaines de reconnaissance de la pluralité de types de dictionnaires de reconnaissance, et les mots de reconnaissance cibles.
 8. Système selon l'une quelconque des revendications précédentes, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole.
 9. Procédé de commande d'un système client-serveur de reconnaissance de la parole pour la reconnaissance par un serveur (200) de la parole fournie en entrée sur un client (100), comprenant :
 une étape d'entrée de parole consistant à fournir de la parole en entrée ; et
 une étape de transmission consistant à transmettre au serveur des données de parole (124b) fournies en entrée lors de l'étape d'entrée de parole, des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) conservé sur le client (100) qui est formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur ;
 une étape de détermination consistant à déterminer un ou plusieurs dictionnaire(s) de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) à partir d'une pluralité de types de dictionnaires de reconnaissance (225) conservés sur le serveur (200), qui sont préparés pour des domaines de reconnaissance respectifs, et le dictionnaire personnel (124a) reçu du client (100) ; et
 une étape de reconnaissance consistant à reconnaître les données de parole (124b) en utilisant au moins le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination.

10. Procédé selon la revendication 9, dans lequel l'étape de reconnaissance comprend une étape consistant à reconnaître les données de parole (124b) en utilisant le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination, et le dictionnaire personnel (124a) reçu du client (100). 5
11. Procédé selon la revendication 9 ou 10, dans lequel l'étape d'entrée de parole comprend une étape d'affichage consistant à afficher une forme d'entrée en tant qu'entrée de parole cible, et dans lequel les informations de gestion de dictionnaires (124c) sont un identificateur de forme d'entrée qui indique un type de forme d'entrée. 10
12. Procédé selon l'une quelconque des revendications 9 à 11, dans lequel les informations de gestion de dictionnaires (124c) contiennent des informations indiquant si le dictionnaire personnel (124a) doit être utilisé dans la reconnaissance des données de parole (124b). 15
13. Procédé selon l'une quelconque des revendications 9 à 12, dans lequel le dictionnaire personnel (124a) est formé en stockant des prononciations et des notations des mots de reconnaissance cibles en correspondance les uns avec les autres. 20
14. Procédé selon la revendication 11, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins un identificateur de forme d'entrée et les mots de reconnaissance cibles en correspondance les uns avec les autres. 25
15. Procédé selon l'une quelconque des revendications 9 à 14, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins l'un des identificateurs de dictionnaires de reconnaissance indiquant des domaines de reconnaissance de la pluralité de types de dictionnaires de reconnaissance, et les mots de reconnaissance cibles. 30
16. Procédé selon l'une quelconque des revendications 9 à 15, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole. 35
17. Mémoire lisible par ordinateur stockant un code de programme de commande d'un système client-serveur de reconnaissance de la parole pour la reconnaissance par un serveur de la parole fournie en entrée sur un client, comprenant : 40
- un code de programme d'une étape d'entrée de parole de la parole d'entrée ; 45
- un code de programme d'une étape de transmission consistant à transmettre au serveur (200) des données de parole (124b) fournies en entrée lors de l'étape d'entrée de parole, des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) conservé sur le client (100), qui est formé en enregistrant des mots de reconnaissance cibles désignés par l'utilisateur ;
- un code de programme d'une étape de détermination consistant à déterminer un ou plusieurs dictionnaire(s) de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) en provenance d'une pluralité de types de dictionnaires de reconnaissance (225) conservés sur le serveur (200), qui sont préparés pour des domaines de reconnaissance respectifs, et le dictionnaire personnel (124a) reçu du client (100) ; et
- un code de programme d'une étape de reconnaissance consistant à reconnaître des données de parole (124b) en utilisant au moins le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination.
18. Serveur de reconnaissance de la parole (200) pour reconnaître de la parole fournie en entrée sur un client (100) et envoyer un résultat de reconnaissance au client (100), comprenant : 50
- un moyen de réception pour recevoir du client (100) des données de parole (124b), des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur ;
- un moyen de conservation de dictionnaires de reconnaissance pour conserver une pluralité de types de dictionnaires de reconnaissance (225) préparés pour des domaines de reconnaissance respectifs ;
- un moyen de détermination (223) pour déterminer un ou plusieurs dictionnaire(s) de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) à partir de la pluralité de types de dictionnaires de reconnaissance (225) et du dictionnaire personnel (124a) reçu du client (100) ; et
- un moyen de reconnaissance (224) pour reconnaître les données de parole en utilisant au moins le dictionnaire de reconnaissance (225) déterminé par ladite étape de détermination (223). 55
19. Serveur selon la revendication 18, dans lequel ledit

- moyen de reconnaissance (224) a pour fonction de reconnaître les données de parole (124b) en utilisant le dictionnaire de reconnaissance (225) déterminé par ledit moyen de détermination (223), et le dictionnaire personnel (124a) reçu du client (100).
20. Serveur selon la revendication 18 ou 19, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole.
21. Client de reconnaissance de la parole (100) pour envoyer de la parole d'entrée devant être reconnue à un serveur (200), et pour recevoir un résultat de reconnaissance de cette parole, comprenant :
- un moyen d'entrée de parole (121, 123) pour fournir en entrée de la parole ;
 - un moyen de conservation de dictionnaire personnel pour conserver un dictionnaire personnel (124a) formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur ; et
 - un moyen de transmission (122) pour transmettre au serveur (200) des données de parole (124b) fournies en entrée par ledit moyen d'entrée de parole, des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et le dictionnaire personnel (124a).
22. Client selon la revendication 21, dans lequel ledit moyen d'entrée de parole (121, 123) comprend un moyen d'affichage (123) pour afficher une forme d'entrée en tant qu'entrée de parole cible, et dans lequel les informations de gestion de dictionnaires (124c) sont un identificateur de forme d'entrée qui indique un type de forme d'entrée.
23. Client selon la revendication 21 ou 22, dans lequel les informations de gestion de dictionnaires (124c) contiennent des informations indiquant si le dictionnaire personnel (124a) doit être utilisé dans la reconnaissance des données de parole.
24. Client selon l'une quelconque des revendications 21 à 23, dans lequel le dictionnaire d'utilisateur (124a) est formé en stockant des prononciations et des notations des mots de reconnaissance cibles en correspondance les uns avec les autres.
25. Client selon la revendication 22, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins un identificateur de forme d'entrée et les mots de reconnaissance cibles en correspondance les uns avec les autres.
26. Client selon l'une quelconque des revendications 21 à 25, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins l'un des identificateurs de dictionnaires de reconnaissance indiquant des domaines de reconnaissance de la pluralité de types de dictionnaires de reconnaissance, et les mots de reconnaissance cibles.
27. Client selon l'une quelconque des revendications 21 à 25, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole.
28. Procédé de commande d'un serveur de reconnaissance de la parole (200) pour reconnaître de la parole fournie en entrée sur un client (100), et envoyer un résultat de reconnaissance au client (100), comprenant :
- une étape de réception consistant à recevoir du client des données de parole (124b), des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur ;
 - une étape de détermination consistant à déterminer un ou plusieurs dictionnaires de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) à partir d'une pluralité de types de dictionnaires de reconnaissance (225) conservés sur le serveur (200), qui sont préparés pour des domaines de reconnaissance respectifs, et du dictionnaire personnel (124a) reçu du client (100) ; et
 - une étape de reconnaissance consistant à reconnaître les données de parole (124b) en utilisant au moins le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination.
29. Procédé selon la revendication 28, dans lequel l'étape de reconnaissance comprend une étape consistant à reconnaître les données de parole (124b) en utilisant le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination, et le dictionnaire personnel (124a) reçu du client (100).
30. Procédé selon la revendication 28 ou 29, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole.
31. Procédé de commande d'un client de reconnaissance de la parole (100) pour envoyer de la parole d'entrée devant être reconnue à un serveur (200) et re-

cevoir un résultat de reconnaissance de cette parole, comprenant :

une étape d'entrée de parole consistant à fournir en entrée de la parole ; et
 une étape de transmission consistant à transmettre au serveur (200) des données de parole (124b) fournies en entrée lors de l'étape d'entrée de parole, des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) conservé sur le client (100), qui est formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur.

32. Procédé selon la revendication 31, dans lequel l'étape d'entrée de parole comprend une étape d'affichage consistant à afficher une forme d'entrée en tant qu'entrée de parole cible, et dans lequel les informations de gestion de dictionnaires (124c) sont un identificateur de forme d'entrée qui indique un type de forme d'entrée.
33. Procédé selon la revendication 31 ou 32, dans lequel les informations de gestion de dictionnaires (124c) contiennent des informations indiquant si le dictionnaire personnel (124a) doit être utilisé dans la reconnaissance des données de parole (124b).
34. Procédé selon l'une quelconque des revendications 31 à 33, dans lequel le dictionnaire personnel (124a) est formé en stockant des prononciations et des notations des mots de reconnaissance cibles en correspondance les uns avec les autres.
35. Procédé selon la revendication 32, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins un identificateur de forme d'entrée et les mots de reconnaissance cibles en correspondance les uns avec les autres.
36. Procédé selon l'une quelconque des revendications 31 à 35, dans lequel le dictionnaire personnel (124a) est formé en stockant également au moins l'un des identificateurs de dictionnaires de reconnaissance indiquant des domaines de reconnaissance de la pluralité de types de dictionnaires de reconnaissance, et les mots de reconnaissance cibles.
37. Procédé selon l'une quelconque des revendications 31 à 36, dans lequel les données de parole (124b) sont des données obtenues en codant ces données de parole.
38. Mémoire lisible par ordinateur qui stocke un code de programme de commande d'un serveur de reconnaissance

de la parole (200) pour la reconnaissance de la parole fournie en entrée sur un client (100) et pour l'envoi d'un résultat de reconnaissance au client (100), comprenant :

un code de programme d'une étape de réception consistant à recevoir du client (100) des données de parole (124b), des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) formé en enregistrant des mots de reconnaissance cibles désignés par un utilisateur ;
 un code de programme d'une étape de détermination consistant à déterminer un ou plusieurs dictionnaire(s) de reconnaissance (225) correspondant aux informations de gestion de dictionnaires (124c) reçues du client (100) à partir d'une pluralité de types de dictionnaires de reconnaissance (225) conservés sur le serveur (200), qui sont préparés pour des domaines de reconnaissance respectifs, et du dictionnaire personnel (124a) reçu du client (100) ; et
 un code de programme d'une étape de reconnaissance consistant à reconnaître les données de parole (124b) en utilisant au moins le dictionnaire de reconnaissance (225) déterminé lors de l'étape de détermination.

39. Mémoire lisible par ordinateur qui stocke un code de programme de commande d'un client de reconnaissance de la parole (100) pour envoyer à un serveur (200) de la parole d'entrée devant être reconnue, et recevoir un résultat de reconnaissance de cette parole, comprenant :

un code de programme d'une étape d'entrée de parole consistant à fournir en entrée de la parole ; et
 un code de programme d'une étape de transmission consistant à transmettre au serveur (200) des données de parole (124b) fournies en entrée lors de l'étape d'entrée de parole, des informations de gestion de dictionnaires (124c) utilisées pour déterminer un domaine de reconnaissance d'un dictionnaire de reconnaissance utilisé pour reconnaître les données de parole, et un dictionnaire personnel (124a) conservé sur le serveur (100), qui est formé en enregistrant des mots de reconnaissance cibles, désignés par un utilisateur.

40. Produit contenant des instructions pouvant être mises en oeuvre sur un processeur et ayant pour effet qu'un dispositif informatique programmable exécute toutes les étapes du procédé selon l'une quelconque

des revendications 28 à 37 lorsque le produit à base d'instructions est exécuté sur ledit dispositif informatique programmable.

5

10

15

20

25

30

35

40

45

50

55

FIG. 1

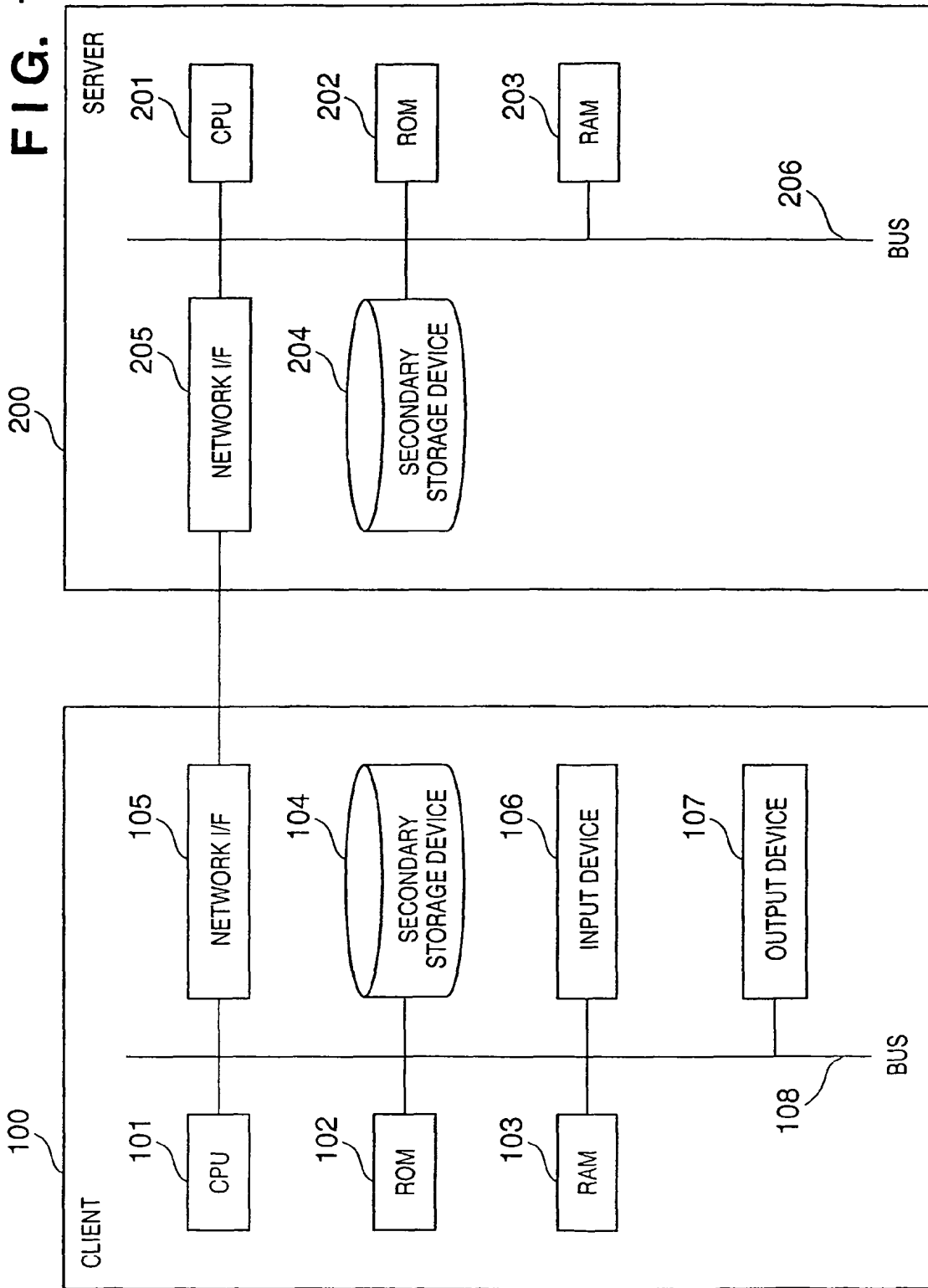


FIG. 2

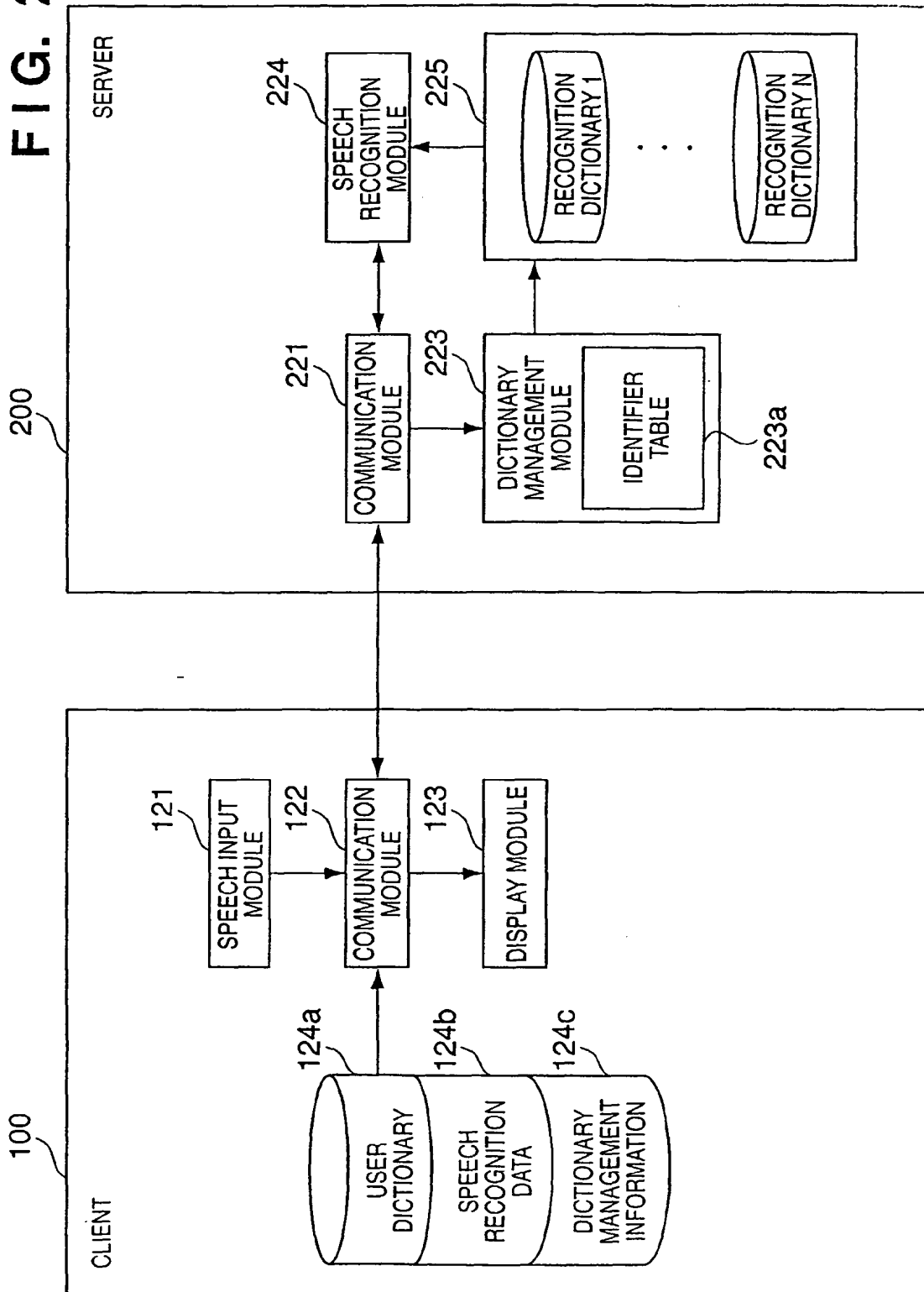


FIG. 3

USER DICTIONARY

PRONUNCIATION	NOTATION
スズキタロー	鈴木太郎
ジタクジュウシヨ	神奈川県川崎市中原区今井上町999
ジタクデンワ	045-222-2222
ケイタイデンワ	090-1111-1111
ファックス	045-333-3333
メール	suzuki@xxx.yyy.co.jp

FIG. 4

NAME	<input type="text"/>
ADDRESS	<input type="text"/>
TEL	<input type="text"/>
FAX	<input type="text"/>
E-MAIL	<input type="text"/>
OTHER	<input type="text"/>

FIG. 5

INPUT FORM IDENTIFIER	RECOGNITION DICTIONARY IDENTIFIER
NAME ADDRESS TEL FAX E-MAIL OTHER	NAME ADDRESS ALPHANUMERIC SYMBOL ALPHANUMERIC SYMBOL ALPHANUMERIC SYMBOL DICTATION

FIG. 6

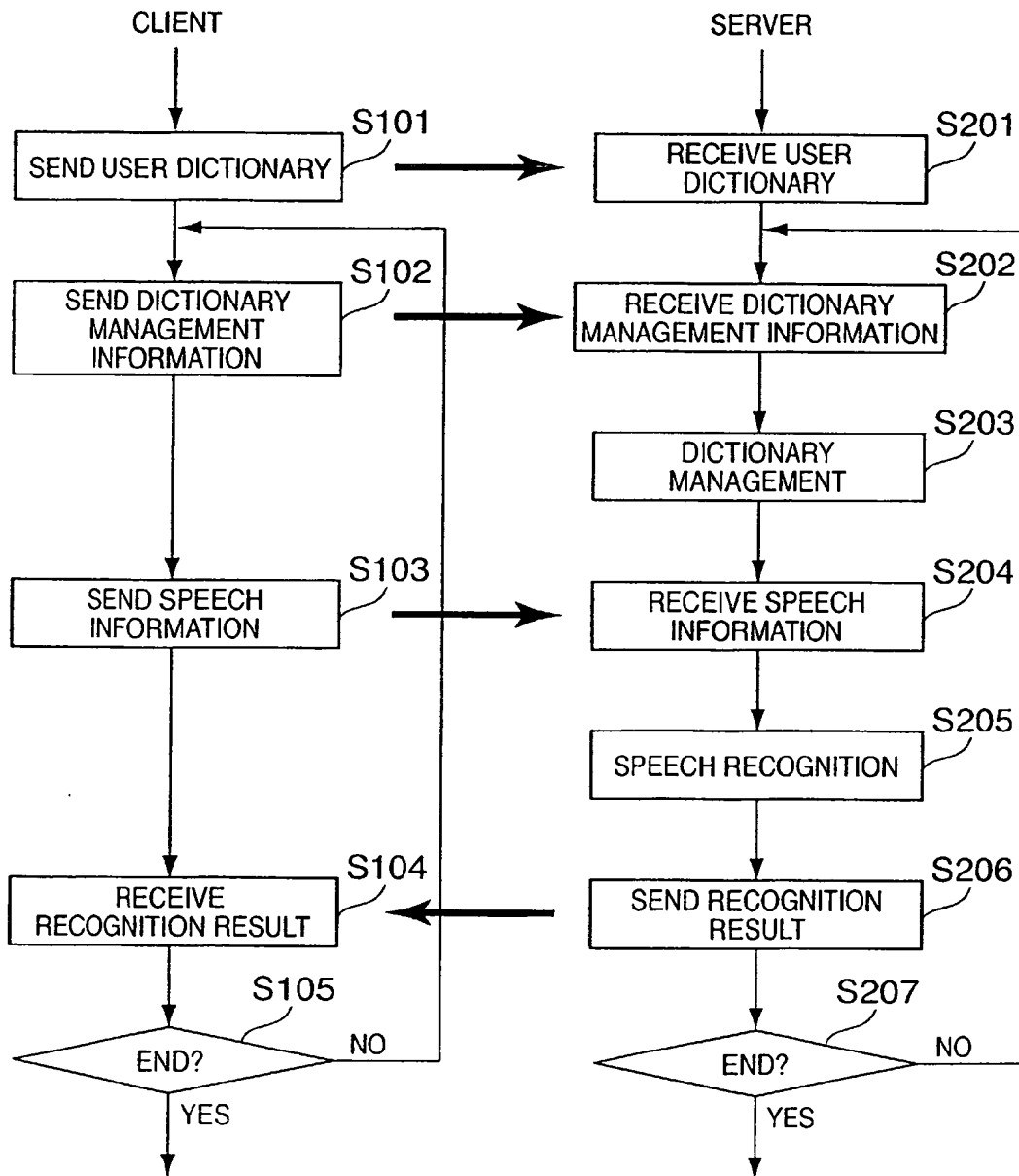


FIG. 7

USER DICTIONARY

PRONUNCIATION	NOTATION	INPUT FORM IDENTIFIER
スズキタロー	鈴木太郎	NAME
ジタクジュウシヨ	神奈川県川崎市中原区今井上町999	ADDRESS
ジタクデンワ	045-222-2222	TEL
ケイタイデンワ	090-1111-1111	TEL
ファックス	045-333-3333	FAX
メール	suzuki@xxx.yyy.co.jp	E-MAIL
スズキタロー	鈴木太郎	OTHER

FIG. 8

USER DICTIONARY

PRONUNCIATION	NOTATION	RECOGNITION DICTIONARY IDENTIFIER
スズキタロー	鈴木太郎	NAME
ジタクジュウシヨ	神奈川県川崎市中原区今井上町53	ADDRESS
ジタクデデンワ	045-222-2222	ALPHANUMERIC SYMBOL
ケイタイデンワ	090-1111-1111	ALPHANUMERIC SYMBOL
ファックス	045-333-3333	ALPHANUMERIC SYMBOL
メール	suzuki@xxx.yyy.co.jp	ALPHANUMERIC SYMBOL
スズキタロー	鈴木太郎	DICTATION